

Sub A' 7.

1. A method of caching in a system for transmitting a plurality of media data titles to one or more client(s) from a central server and a proxy server, wherein each title is divided into blocks to be transmitted to the one or more client(s) in a time sequence, and each block is divided into sub-blocks, comprising:
- identifying which sub-blocks from different blocks of each title that are to be cached, wherein the identified sub-blocks include ones that are distributed over the blocks of at least one title; and
- caching the identified sub-blocks under the control of the proxy server to reduce the transmission bit rate of the central server for transmitting the titles.
2. The method of claim 1, wherein the cached sub-blocks are cached for time periods that are independent of passage of time.
3. The method of claim 1, wherein the caching caches substantially the same number of sub-blocks from each block of said at least one title.
4. The method of claim 1, wherein the media titles include video titles, and the sub-blocks include video frames, and each block is divided into video frames that are to be transmitted sequentially, and further comprising inserting the cached video frames into a stream of video frames from the central server to form a combined stream and sending the combined stream to the client(s).
5. The method of claim 1, wherein the media titles include video titles, and the sub-blocks comprise partial information of video frames, wherein the video frames are to be transmitted sequentially, further comprising combining the partial information of video frames from the proxy server with complementary partial information of such video frames from the central server into complete video frames and sending the complete video frames to the client(s).

5

10

15

20

25

30

**THE UNIVERSITY OF CHICAGO**

Sub A' 7

5            13. The method of claim 1, further comprising beginning a caching process at the proxy server after receiving a title request from a client by ensuring there is sufficient bandwidth from said proxy to such client to deliver the request and if not, denying the request.

15. The method of claim 13, further comprising:  
15 when said proxy server does not have complete content of the requested title,  
determining if there is sufficient available backbone bandwidth to carry said title  
from the central server to said proxy server and if not, rejecting the request.

17. The method of claim 1, further comprising replacing a cached portion of a particular title by deleting the most recently cached portion of such title.

18. The method of claim 1, further comprising deciding which titles shall be subject to caching replacement using a most current access profile as an indication of a future profile.

[illegible]

Sub A' 7

20. The method of claim 1, further comprising deciding which titles shall be subject to caching replacement using a current access profile as an indication of the future profile, wherein said deciding includes:

calculating an access frequency of each title in a storage of the proxy server, said access frequency being a function of the accesses to such title during the time window or a portion thereof; and

21. The method of claim 20, wherein said access frequency is proportional to the sum of the accesses to such title during the time window or a portion thereof.

23. The method of claim 1, further comprising detecting multiple ongoing requests from clients for a title received at different times during caching in response to an initial request of said title, and increasing the number of sub-blocks cached from the blocks of at least one title in response to a subsequent request of said title.

24. A method of deciding an amount of a data title for a caching action comprising:

**090806Z**

Sub A' 7

defining a parameter  $N$  to specify caching units; and  
each time a caching action is activated, performing said action on one or more  $1/N$  portions of said data title (where  $N \geq 1$ ), wherein said  $1/N$  portion of said data title is distributed in said data title.

5

25. The method of claim 24, wherein said  $1/N$  portion of said data title is evenly distributed by:

dividing said data title into a number of blocks;

dividing each block into a number of sub-blocks; and

10

for each caching action, selecting the same number of sub-block(s) from each block of at least one title.

15

26. The method of claim 24, wherein said caching action comprises adding data to a cached portion of said title.

27. The method of claim 24, wherein said caching action comprises removing data from a cached portion of said title.

20

28. A system for delivering media information; the system comprising:  
a plurality of proxy servers, each servicing a number of terminal devices and receiving a request from one of said terminal devices when a user of said one of said terminal devices desires for a media title; each of said proxy servers comprising a cache memory for storing units of some of titles; wherein the number of units of each of said titles is determined by a request frequency to said each of said titles; and

25

a central media server coupled to said proxy servers; said central media server having a storage space for storing a plurality of said titles and providing data from one of said titles when receiving a proxy request from one of said proxy servers.

30

29. A system for delivering media information; the system comprising:

Sub A<sup>c</sup> 7

a plurality of proxy servers, each servicing a number of terminal devices and receiving a request from one of said terminal devices when a user of said one of said terminal devices desires a media title from a plurality of media titles; wherein at least one of said proxy servers comprises a cache memory storing a number of units of at least one of said titles; wherein the units of the at least one title stored are distributed over such title; and

a central server coupled to said proxy servers; said central server having a storage space for storing a plurality of said titles and providing data from one of said titles when receiving a proxy request from one of said proxy servers.

30. The system of claim 29, wherein said at least one proxy server stores the units of the at least one title for time periods that are independent of passage of time.

31. The system of claim 29, wherein the at least one title is divided into blocks to be transmitted to the one or more user(s) in a time sequence, and each block is divided into sub-blocks, wherein the at least one proxy server caches the same number of sub-blocks from each block of said at least one title.

32. The method of claim 29, wherein the at least one title includes a video divided into blocks to be transmitted in a time sequence, and each block is divided into sub-blocks, and the sub-blocks comprise partial information of video frames, wherein the video frames are to be transmitted sequentially.

33. The system of claim 32, wherein the at least one proxy server combines the partial information of video frames with complementary partial information of such video frames from the central server into complete video frames and sends the complete video frames to user(s)

Sub A 7

34. The system of claim 33, wherein the partial information comprise video information along some scan lines of video frames, and wherein the at least one proxy server combines the video information along such scan lines with complementary video information along other scan lines of such video frames from the central server into complete video frames and sends the complete video frames to the user(s).

35. The system of claim 33 wherein the partial information comprise video information obtained at a set of sampling times and at a first sampling rate lower than that of a video source from which said information originates, and wherein the at least one proxy server combines the video information at the lower first sampling rate with complementary video information taken at sampling times different from the set of sampling times of such video frames from the central server into video data at a sampling rate higher than the first sampling rate and sends the video data at the higher sampling rate to the user(s).

36. The system of claim 32, wherein the sub-blocks comprise information in a base layer of a scalable multilayer system.

37. The system of claim 29, where the number of units is a function of an access profile of the at least one title at the at least one proxy server.

38. A method of caching in a system for transmitting a plurality of data titles to one or more client(s) from a central server and a proxy server, wherein each title is divided into blocks to be transmitted to the one or more client(s) in a time sequence, and each block is divided into sub-blocks, comprising:

identifying which sub-blocks from different blocks of each title that are to be cached, wherein the identified sub-blocks include ones that are distributed over the blocks of at least one title; and

0966705-090800

5

10

15

20

20

25

25

30



Sub A' 7

**SECRET**

45. The device of claim 41, wherein the media titles include video titles,  
5 and the sub-blocks comprise partial information of video frames, wherein the video  
frames are to be transmitted sequentially, said method further comprising combining  
the partial information of video frames from the proxy server with complementary  
partial information of such video frames from the central server into complete video  
frames and sending the complete video frames to the client(s).

46. The device of claim 45, wherein the partial information comprise video information along some scan lines of video frames, said method further comprising combining the video information along such scan lines with complementary video information along other scan lines of such video frames from the central server into complete video frames and sending the complete video frames to the client(s).

47. The device of claim 45, wherein the partial information comprise video information obtained at a set of sampling times and at a first sampling rate lower than that of a video source from which said information originates, said method further comprising combining the video information at the lower first sampling rate from the proxy server with complementary video information taken at sampling times different from the set of sampling times of such video frames from the central server into video data at a sampling rate higher than the first sampling rate and sending the video data at the higher sampling rate to the client(s).

48. The device of claim 45, wherein the partial information comprise video information in a basic layer and the complementary partial information comprises video information in an enhancement layer, said basic and enhancement layers being defined according to spatial, signal-to-noise or temporal scalability.

Sub A 7  
49. The device of claim 41, wherein the identifying is made as a function of an access profile of the titles at the proxy.

50. The device of claim 41, wherein prior to any accesses of the titles by the client(s), an average caching approach utilizes storage at the proxy server by storing a substantially equal number of sub-blocks from each title.

51. The device of claim 41, wherein prior to any accesses of the titles by the client(s), a proportional caching approach utilizes access history data to determine how much of each title to cache.

52. The device of claim 41, wherein after the system starts operation, cache content at the proxy server will change from time to time to reflect actual access behavior.

53. The device of claim 41, said method further comprising beginning a caching process at the proxy server after receiving a title request from a client by ensuring there is sufficient bandwidth from said proxy to such client to deliver the request and if not, denying the request.

54. The device of claim 53, said method further comprising delivering the complete content of the requested title when such content is in local storage of said proxy server.

55. The device of claim 53, said method further comprising:  
when said proxy server does not have complete content of the requested title, determining if there is sufficient available backbone bandwidth to carry said title from the central server to said proxy server and if not, rejecting the request.

56. The device of claim 55, said method further comprising activating a progressive caching process to adjust cache content at said proxy server to reflect the requested title.

5 57. The device of claim 41, said method further comprising replacing a cached portion of a particular title by deleting the most recently cached portion of such title.

10 58. The device of claim 41, said method further comprising deciding which titles shall be subject to caching replacement using a most current access profile as an indication of a future profile.

15 59. The device of claim 41, said method further comprising keeping track of each access request at the proxy server in order to determine which titles shall be subject to caching replacement.

20 60. The device of claim 41, said method further comprising deciding which titles shall be subject to caching replacement using a current access profile as an indication of the future profile, wherein said deciding includes:

25 defining a time window ending at the time of the caching replacement;  
calculating an access frequency of each title in a storage of the proxy server, said access frequency being a function of the accesses to such title during the time window or a portion thereof; and  
performing the caching replacement in response to the access frequencies of the titles in the storage.

61. The device of claim 60, wherein said access frequency is proportional to the sum of the accesses to such title during the time window or a portion thereof.

008060" 50/85960

Sub A' 7

Sub A' 7

62. The device of claim 60, wherein said access frequency is proportional to a time-weighted sum of the accesses to such title during the time window or a portion thereof, with the time weighting in favor of accesses occurring more recently in the window.

5

63. The device of claim 41, said method further comprising detecting multiple ongoing requests from clients for a title received at different times during caching in response to an initial request of said title, and increasing the number of sub-blocks cached from the blocks of at least one title in response to a subsequent request of said title.

10

64. A computer readable storage device embodying a program of instructions executable by a computer to perform a method of deciding an amount of a data title for a caching action, said method comprising:

15

defining a parameter  $N$  to specify caching units; and

each time a caching action is activated, performing said action on one or more  $1/N$  portions of said data title (where  $N \geq 1$ ), wherein said  $1/N$  portion of said data title is distributed in said data title.

20

65. The device of claim 64, wherein said  $1/N$  portion of said data title is evenly distributed by:

dividing said data title into a number of blocks;

dividing each block into a number of sub-blocks; and

25

for each caching action, selecting the same number of sub-block(s) from each block of at least one title.

66. The device of claim 64, wherein said caching action comprises adding data to a cached portion of said title.

09658705-090800

Sub A<sup>7</sup>

67. The device of claim 64, wherein said caching action comprises removing data from a cached portion of said title.

68. A method for transmitting a program of instructions executable by a computer to perform a process of caching in a system for transmitting a plurality of media data titles to one or more client(s) from a central server and a proxy server, wherein each title is divided into blocks to be transmitted to the one or more client(s) in a time sequence, and each block is divided into sub-blocks, said method comprising:

transmitting to a client device a program of instructions; and  
enabling the client device to perform, by means of such program, the following process:

identifying which sub-blocks from different blocks of each title that are to be cached, wherein the identified sub-blocks include ones that are distributed over the blocks of at least one title; and

caching the identified sub-blocks under the control of the proxy server to reduce the transmission bit rate of the central server for transmitting the titles.

69. The method of claim 68, wherein the program enables the cached sub-blocks to be stored for time periods that are independent of passage of time.

70. The method of claim 68, wherein the program enables the caching to cache substantially the same number of sub-blocks from each block of said at least one title.

71. The device of claim 68, wherein the media titles include video titles, and the sub-blocks comprise partial information of video frames, wherein the video frames are to be transmitted sequentially, wherein the program enables the client device to further perform, by means of such program, the following: combining the partial information of video frames from the proxy server with complementary partial

information of such video frames from the central server into complete video frames and sending the complete video frames to the client(s).

72. The method of claim 71, wherein the partial information comprise  
5 video information in a basic layer and the complementary partial information  
comprises video information in an enhancement layer, said basic and enhancement  
layers being defined according to spatial, signal-to-noise or temporal scalability.

73. The method of claim 68, wherein the identifying is made as a function  
10 of an access profile of the titles at the proxy.

74. A method for transmitting a program of instructions executable by a computer to perform a process of deciding an amount of a data title for a caching action, said method comprising:

15 transmitting to a client device a program of instructions; and  
enabling the client device to perform, by means of such program, the  
following process:

defining a parameter  $N$  to specify caching units; and  
each time a caching action is activated, performing said action on one or  
20 more  $1/N$  portions of said data title (where  $N \geq 1$ ), wherein said  $1/N$  portion of said  
data title is distributed in said data title.

75. The method of claim 74, wherein said  $1/N$  portion of said data title is evenly distributed by:

25       dividing said data title into a number of blocks;  
dividing each block into a number of sub-blocks; and  
for each caching action, selecting the same number of sub-block(s) from each  
block of at least one title.

**SECRET**

Sub A<sup>1</sup> 7

49

76. The device of claim 74, wherein said caching action comprises adding data to a cached portion of said title.

77. The device of claim 74, wherein said caching action comprises  
5 removing data from a cached portion of said title.

Add A<sup>2</sup> 7

008060" 50785960